

P20721.P06



#3

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Xiaoning NIE

Serial No. : 09/817,031

Group Art Unit: 2152

Filed : March 27, 2001

Examiner : Unknown

For : SERVER MODULE FOR MODULARLY DESIGNED SERVER

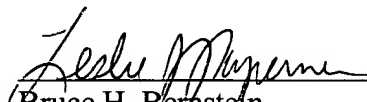
CLAIM OF PRIORITY

Commissioner of Patents and Trademarks
Washington, D.C. 20231

Sir:

Applicant hereby claims the right of priority granted pursuant to 35 U.S.C. 119 based upon German Application No. 100 16 236.3, filed March 31, 2000. As required by 37 C.F.R. 1.55, a certified copy of the German application is being submitted herewith.

Respectfully submitted,
Xiaoning NIE


Bruce H. Bernstein
Reg. No. 29,027

June 11, 2001
GREENBLUM & BERNSTEIN, P.L.C.
1941 Roland Clarke Place
Reston, VA 20191
(703) 716-1191

BUNDESREPUBLIK DEUTSCHLAND



Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

**CERTIFIED COPY OF
PRIORITY DOCUMENT**

Aktenzeichen:

100 16 236.3

Anmeldetag:

31. März 2000

Anmelder/Inhaber:

Infineon Technologies AG, München/DE

Bezeichnung:

Servermodul für modular aufgebauten Server

IPC:

G 06 F, H 04 L

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 22. März 2001
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

11/10/01

Beschreibung

Servermodul für modular aufgebauten Server

- 5 Die Erfindung betrifft ein Servermodul für einen modular aufgebauten Server.

Server sind Berechnungseinrichtungen, die in einem verteilten Rechnersystem zentrale Dienste für mehrere Nutzer bzw.

- 10 Clients anbieten. Die Server und Nutzer sind zumeist über ein LAN miteinander verbunden. Server können entsprechend ihrer Funktion beispielsweise in File-Server, Anwendungsserver, Print-Server, CD-ROM-Server, E-Mail-Server unterschieden werden. Ein File-Server verwaltet zentral Daten und/oder Dateien
15 und stellt diese anderen Rechnern zur Verfügung. Anwendungsserver stellen dem Netzwerk oder den Anwendern zentral Programme zur Verfügung. Ein Print-Server bzw. Druckserver verwaltet zentral Druckaufträge und verteilt sie an die entsprechenden Drucker im Netz. Ein CD-ROM-Server verfügt über mehrere
20 CD-ROM-Laufwerke, auf die über das LAN zugegriffen werden kann. Ein E-Mail-Server dient für alle Teilnehmer als Mailbox und verwaltet alle ein- und abgehenden E-Mails.

- 25 Fig. 1 zeigt den Aufbau eines Netzwerkes mit mehreren Servern nach dem Stand der Technik. Die Server A-D übernehmen verschiedene Funktionen innerhalb des Netzwerkes, beispielsweise ist Server A ein File-Server, Server B ein Anwendungsserver, Server C ein Print-Server, Server D ein CD-ROM-Server. Die Server sind über eine Kommunikationsschnittstelle mit einem
30 LAN und über eine Router-Schnittstelle mit einem Router zum Datenaustausch mit den übrigen Servern verbunden. Der Datenaustausch erfolgt mittels Datenpaketen. Der Router ist ein Verknüpfungsrechner zwischen den beiden LAN-Netzen und stellt einen Teil der untersten drei Schichten des OSI-
35 Referenzmodells dar. Als Datenpaketvermittler wertet der Router die in den Datenpaketen enthaltenen Schicht-3-Informationen, beispielsweise die IP-Adressen aus und ermit-

telt anhand von Routing-Tabellen den günstigsten weiteren Datenweg. Der Router kann dabei auch Datenpakete über ein LAN bzw. über Telekommunikationsnetze senden oder empfangen.

- 5 Die mit dem Router verbundenen Server nach dem Stand der Technik weisen zwei Hauptbestandteile auf, nämlich eine Datenverarbeitungseinheit bzw. einen Prozessor sowie eine Netzwerk-Schnittstellenkarte NIC. Falls der an das LAN1 angeschlossene Server A mit dem an dem LAN2 angeschlossenen Server B kommunizieren möchte, müssen die Daten in das geeignete Rahmenformat gebracht werden. In der Netzwerk-Schnittstellenkarte NIC_a des Servers A werden die an den Server B zu übertragenden Daten in das Datenformat des Netzwerks LAN1 eingepackt (Framing) und an die Line-Card des Routers gesendet.
- 10
- 15 Durch die Line-Card des Routers werden die Daten ausgepackt (Deframing) und dann in das Datenformat des zweiten Netzwerkes LAN2 gebracht bzw. eingepackt (Framing). Anschließend sendet der Router die umformatierten Daten an den Server B. Die Datenpakete, die der Server A an den Router absendet,
- 20
- enthalten als Zieladresse den Server B.

Ein Nachteil der in Fig. 1 dargestellten Netzwerkanordnung besteht darin, daß der technische Aufwand zur Übertragung von Daten von einem Server zu einem anderen Server über den Router technisch relativ aufwendig ist, da ein ständiges Ein- und Auspacken bzw. eine Datenformatumwandlung erfolgen muß. Die dazu in den Servern vorhandenen Netzwerk-Schnittstellenkarten besitzen eine hohe schaltungstechnische Komplexität, so daß die Server relativ hohe Herstellungskosten aufweisen.

30

Ein weiterer Nachteil der in Fig. 1 dargestellten Netzwerkanordnung besteht darin, daß stets ein Router zwischen den Servern vorhanden sein muß. Die Anzahl der möglichen Server in der in Fig. 1 gezeigten Netzwerkanordnung wird daher durch die Leistungsfähigkeit und die Anzahl der Server-Anschlüsse des Routers begrenzt. Eine flexible Erweiterung der in Fig. 1

35

dargestellten herkömmlichen Netzwerkanordnung wird hierdurch erschwert.

5 Es ist daher die Aufgabe der vorliegenden Erfindung, ein Servermodul für einen modular aufgebauten Server zu schaffen, der einen Datenaustausch mit einem geringen schaltungstechnischen Aufwand ermöglicht.

10 Diese Aufgabe wird erfindungsgemäß durch ein Servermodul mit den im Patentanspruch 1 angegebenen Merkmalen gelöst.

Die Erfindung schafft ein Servermodul für einen modular aufgebauten Server mit mindestens einer Datenverarbeitungseinheit zur Datenverarbeitung von Datenpaketen,
15 mindestens einer adressierbaren Kommunikationsschnittstelle zum Anschluß des Servermoduls an ein externes Netzwerk, über das die Datenpakete übertragen werden,
einer Schaltschnittstelle zum Anschluß des Servermoduls an eine Schalteinrichtung des modular aufgebauten Servers, und
20 mit
einer Routing-Berechnungseinheit zur Berechnung einer Servermodul-Adresse mittels einer Routing-Tabelle in Abhängigkeit von der Auslastung der Datenverarbeitungseinheiten aller Servermodule des modular aufgebauten Servers.

25

Bei einer bevorzugten Ausführungsform ist das Servermodul zur Datenverarbeitung von Datenpaketen eines bestimmten vorgegebenen Applikationstyps vorgesehen.

30 Dabei wird die Servermodul-Adresse vorzugsweise in Abhängigkeit von dem Applikationstyp des übertragenen Datenpakets berechnet.

Die Kommunikationsschnittstellen weisen vorzugsweise Pufferspeicher zum Zwischenspeichern von übertragenen Datenpaketen auf.
35

Die Routing-Tabelle wird bei einer besonders bevorzugten Ausführungsform in einem zugewiesenen Routing-Servermodul des modular aufgebauten Servers ständig aktualisiert.

- 5 Das Routing-Servermodul überträgt die aktuelle Routing-Tabelle an die übrigen Servermodule vorzugsweise über die Schaltschnittstelle.

- 10 Bei einer weiteren bevorzugten Ausführungsform sammelt das Routing-Servermodul Daten über die Auslastung der Datenverarbeitungseinheiten aller Servermodule des modular aufgebauten Servers und wertet diese aus.

- 15 Dabei aktualisiert das Routing-Servermodul die Routing-Tabelle in Abhängigkeit von den ausgewerteten Auslastungsdaten, den zugewiesenen Applikationstypen der übrigen Servermodule sowie Prioritätsinformationsdaten der übertragenen Datenpakete.

- 20 Bei einer weiteren bevorzugten Ausführungsform werden bei einer zu hohen Auslastung der Datenverarbeitungseinheiten eines Servermoduls die in diesem Servermodul ablaufenden Datenverarbeitungsprozesse zumindestens teilweise auf Datenverarbeitungseinheiten anderer Servermodule übertragen.

25

Im weiteren wird eine bevorzugte Ausführungsform des erfindungsgemäßen Servermoduls für einen modular aufgebauten Server unter Bezugnahme auf die beigefügten Figuren zur Erläuterung erfindungswesentlicher Merkmale beschrieben.

30

Es zeigen:

Fig. 1 eine Netzwerkanordnung nach dem Stand der Technik;

- 35 Fig. 2 einen aus mehreren Servermodulen modular aufgebauten Server gemäß der Erfindung;

Fig. 3 ein Blockschaltbild eines erfindungsgemäßen Servermoduls.

In Fig. 2 ist ein modular aufgebauter Server 1 dargestellt,
5 der modular aus mehreren verschiedenen Servermodulen 2, 3, 4, 5 aufgebaut ist. Jedem Servermodul 2, 3, 4, 5 ist zur Datenverarbeitung von Datenpaketen vorgesehen, die einem bestimmten Applikationstyp entsprechen. Beispielsweise sind die Servermodule 3, 5 File-Server zur Datenverarbeitung von FTP-
10 Dateien, während das Servermodul 4 einen Druckserver darstellt. Das Servermodul 2 ist über einen internen Datenbus bzw. eine Datenleitung 6 mit einer Schalteinrichtung 7 verbunden. Die Schalteinrichtung 7 ist über weitere interne Datenbusse bzw. Datenleitungen 8, 9, 10 mit den übrigen Server-
15 modulen 3, 4, 5 des modular aufgebauten Servers 1 verbunden. Die Servermodule 2, 4 sind über Anschluß-Datenbusse bzw. Datenleitungen 11, 12 an ein erstes lokales Netzwerk 13 angeschlossen und die Servermodule 3, 5 sind über Anschluß-Datenbusse 14, 15 an ein zweites lokales Netzwerk 16 ange-
20 schlossen.

Fig. 3 ist ein Blockschaltbild zur Erläuterung des Aufbaus eines erfindungsgemäßen Servermoduls des modular aufgebauten Servers 1. Alle Servermodule 2, 3, 4, 5 des modular aufgebauten Servers 1 weisen den gleichen schaltungstechnischen Aufbau auf. Der Aufbau eines Servermoduls wird in Fig. 3 anhand
25 des Servermoduls 2 dargestellt. Das Servermodul 2 enthält mindestens eine Datenverarbeitungseinheit 17, eine Routing-Berechnungseinheit 18, mindestens eine adressierbare Kommunikationsschnittstelle 19 sowie eine Schaltschnittstelle 20.
30 Die Datenverarbeitungseinheit 17 kann einen oder mehrere Prozessoren zur Datenverarbeitung von Datenpaketen enthalten. Über die adressierbare Kommunikationsschnittstelle und den Anschluß-Datenbus 11 ist das Servermodul 2 mit dem externen
35 Netzwerk 13 verbunden. Die von dem Netzwerk 13 empfangenen externen Datenpakete gelangen über den Anschluß-Datenbus 11 zu der Kommunikationsschnittstelle 19. Das ankommende Daten-

paket ist ein Daten-Rahmen, der aus Kopfdaten (Header) und Nutzdaten (Payload) besteht. In den Kopfdaten ist eine Leitungsschichtadresse bzw. Layer 2-Adresse enthalten. Darüber hinaus enthält der ankommende Internet-Datenrahmen Informationen über den Applikationstyp. Die Routing-Berechnungseinheit 18 prüft, ob die Leitungsschicht-Adresse des ankommenden Datenpakets ein Servermodul innerhalb des modular aufgebauten Servers 1 adressiert oder nicht. Falls der modular aufgebaute Server 1 nicht adressiert wird, liest die Routing-Berechnungseinheit aus einer Routing-Tabelle die Port-Adresse des externen Zielrechners aus und versendet das Datenpaket über die Kommunikationsschnittstelle 19 an diesem externen Zielrechner. Falls umgekehrt durch die in dem ankommenden Datenpaket enthaltene Leitungsschichtadresse bzw. durch die MAC-Informationsdaten ein Servermodul innerhalb des modular aufgebauten Servers 1 adressiert wird, wird zunächst der Applikationstyp des empfangenen Datenrahmens bestimmt. Aus der aktuellen Routing-Tabelle wird in Abhängigkeit von der Rechnerauslastung der Datenverarbeitungseinheiten innerhalb der verschiedenen Servermodule 2, 3, 4, 5 des modular aufgebauten Servers 1 die Servermodul-Adresse desjenigen Servermoduls ausgelesen, das diesem Applikationstyp zugeordnet ist und deren Datenverarbeitungseinheiten am wenigsten ausgelastet sind. An diese ausgelesene Servermodul-Adresse versendet die Routing-Berechnungseinheit die empfangenen Nutzdaten als ein Internet-Protokoll-Datenpaket bzw. IP-Datagramm über die internen Datenbusleitungen und die Schalteinrichtung 7. Hierzu generiert die Routing-Berechnungseinheit 18 entweder einen internen Header mit der ermittelten Servermodul-Adresse, der an das IP-Datagramm angefügt wird, oder steuert die Schalteinrichtung 7 mittels (nicht dargestellter) Steuerleitungen derart an, daß das aus den Nutzdaten des ankommenden Datenpakets gewonnene Internet-Protokoll-Datenpaket an dasjenige Servermodul geschaltet wird, welches durch die aus der Routing-Tabelle stammende Servermodul-Adresse adressiert wird.

Bei dem modular aufgebauten Server 1 übernehmen die verschiedenen Servermodule 2, 3, 4, 5 verschiedenen Funktionen. Ein Servermodul wird als Routing-Tabellenmanager mit der Aktualisierung der Routing-Tabelle betraut. Ein weiteres Servermodul
5 bestimmt als Load-Manager die Rechnerauslastung aller Datenverarbeitungseinheiten der verschiedenen Servermodule innerhalb des modular aufgebauten Servers 1. Das mit der Lastermittlung betraute Servermodul sammelt ständig Daten über die Rechnerauslastung der Datenverarbeitungseinheiten der übrigen
10 Servermodule und wertet diese aus. Die Auswertungsergebnisse werden an dasjenige Servermodul übertragen, das die Routing-Tabelle verwaltet. Die Routing-Tabelle wird dort in Abhängigkeit von den übertragenen Auslastungsdaten, den zugewiesenen Applikationstypen der Servermodule sowie von Prioritätsinfor-
15 mationsdaten der ankommenden Datenpakete ständig aktualisiert. Datenkopien der aktualisierten Routing-Tabelle werden an die übrigen Servermodule versendet, so daß jedes Servermodul 2-5 innerhalb des modular aufgebauten Servers 1 über den Auslastungszustand der übrigen Servermodule informiert ist.
20 Bei einer zu hohen Auslastung der Datenverarbeitungseinheiten eines Servermoduls werden dabei vorzugsweise Datenverarbeitungsprozesse, die in diesem Servermodul ablaufen, zumindestens teilweise auf Datenverarbeitungseinheiten von anderen Servermodulen übertragen. Die Rechnerauslastung der verschiedenen Servermodule werden hierdurch weitgehend aneinander an-
25 geglichen.

Sind beispielsweise die Servermodule 3, 5 als File-Server zur Bearbeitung von FTP-Applikationen geeignet und erhält das
30 Servermodul 2 über den Anschluß 11 von dem lokalen Netzwerk 13 ein externes Datenpaket mit einer FTP-Applikation, bestimmt es anhand der in der Routing-Berechnungseinheit 18 enthaltenen Routing-Tabelle die Servermodul-Adresse desjenigen File-Servers, der die geringste Rechnerauslastung auf-
35 weist. Hierzu befindet sich eine Kopie der aktualisierten Routing-Tabelle in einer Speichereinrichtung der Routing-Berechnungseinheit 18. Ist beispielsweise die Rechnerausla-

stung des als File-Server arbeitenden Servermoduls 3 geringer als die Rechnerauslastung des ebenfalls als File-Server arbeitenden Servermoduls 5, ist die Routing-Tabelle durch den Routing-Tabellenmanager derart aktualisiert, daß bei einem ankommenden Datenpaket, welches einen File-Server benötigt, als Servermodul-Adresse die Adresse des Servermoduls 3 durch die Routing-Berechnungseinheit 18 generiert wird. Das Internet-Protokoll-Datenpaket, das in dem Internet-Datenpaket als Payload bzw. Nutzdaten enthalten ist, wird durch die Routing-Berechnungseinheit 18 vorzugsweise mit einem internen Header bzw. Kopfdaten versehen, die als Servermodul-Adresse die Adresse des Servermoduls 3 beinhalten. Die mit dem internen Header versehenen Internet-Protokoll-Datenpakete werden als interne Datenpakete über den internen Datenbus 6, die Schalteinrichtung 7 und den internen Datenbus 8 an das Servermodul 3 übertragen. Hierzu liest die Schalteinrichtung 7 die in dem Header des internen Datenpakets enthaltene interne Servermodul-Adresse aus und leitet das daran hängende Internet-Protokoll-Datenpaket an das adressierte Servermodul 3 weiter.

Das erfindungsgemäße Servermodul besitzt eine integrierte Router-Funktion, so daß der modular aufgebaute Server 1 keinen eigenständigen Router benötigt. Der in Fig. 2 dargestellte modular aufgebaute erfindungsgemäße Server 1 besitzt eine Schalteinrichtung 7, an die im Prinzip eine beliebig hohe Anzahl von Servermodulen 2-5 anschließbar sind. Eine Erweiterung des modular aufgebauten Servers 1 mit weiteren Servermodulen zur Leistungssteigerung und zur Übernahme zusätzlicher Serverdienste ist daher einfach durch Anschluß zusätzlicher Servermodule an die Schalteinrichtung 7 flexibel möglich. Da jedes Servermodul über eine eigene Routing-Berechnungseinheit verfügt, wird der Lastausgleich zwischen den verschiedenen Servermodulen innerhalb des modular aufgebauten Servers 1 optimiert. Jedes Servermodul leitet ankommende Datenpakete an dasjenige Servermodul weiter, das für den Applikationstyp zuständig ist und das die geringste Rechnerauslastung aufweist. Ferner werden ablaufende Datenverarbeitungsprozesse, die in-

nerhalb der Datenverarbeitungseinheiten eines überlasteten Servermoduls ablaufen, an weniger belastete Servermodule übergeben.

- 5 Der modulare Aufbau des Servers 1 gewährleistet somit eine optimal ausgeglichene Rechnerauslastung sowie flexible Erweiterungsmöglichkeiten. Ein umständliches Umformatieren von externen Datenpaketformaten in interne Datenpaketformate entfällt, da die in den externen Datenpaketen bzw. Internet-
- 10 Rahmen enthaltenen IP-Datenpakete unverändert an die entsprechenden Servermodule weitergeschaltet werden. Der schaltungs-technische Aufwand zur internen Weiterleitung empfangener Datenpakete an die übrigen Servermodule ist daher gering.

Patentansprüche

1. Servermodul für modular aufgebauten Server (1) mit
mindestens einer Datenverarbeitungseinheit (17) zur Datenver-
5 arbeitung von Datenpaketen;
mindestens einer adressierbaren Kommunikationsschnittstelle
(19) zum Anschluß des Servermoduls (2, 3, 4, 5) an ein exter-
nes Netzwerk (13), über das die Datenpakete übertragen wer-
den;
10 einer Schaltschnittstelle (20) zum Anschluß des Servermoduls
(2, 3, 4, 5) an eine Schalteinrichtung (7) des modular aufge-
bauten Servers (1); und mit
einer Routing-Berechnungseinheit (18) zur Berechnung einer
Servermodul-Adresse mittels einer Routing-Tabelle in Abhän-
15 gigkeit von der Auslastung der Datenverarbeitungseinheiten
aller Servermodule (2, 3, 4, 5) des modular aufgebauten
Servers (1).
2. Servermodul nach Anspruch 1,
20 d a d u r c h g e k e n n z e i c h n e t,
daß das Servermodul (2, 3, 4, 5) zur Datenverarbeitung von
Datenpaketen bestimmter vorgegebener Applikationstypen vorge-
sehen ist.
- 25 3. Servermodul nach Anspruch 1 oder 2,
d a d u r c h g e k e n n z e i c h n e t,
daß die Servermodul-Adresse in Abhängigkeit von dem Applika-
tionstyp des übertragenen Datenpakets berechnet wird.
- 30 4. Servermodul nach einem der vorangehenden Ansprüche,
d a d u r c h g e k e n n z e i c h n e t,
daß die Kommunikationsschnittstelle (19) Pufferspeicher zum
Zwischenspeichern von übertragenen Datenpaketen aufweist.
- 35 5. Servermodul nach einem der vorangehenden Ansprüche,
d a d u r c h g e k e n n z e i c h n e t,

daß die Routing-Tabelle in einem zugewiesenen Routing-Servermodul des modular aufgebauten Servers (1) ständig aktualisiert wird.

5 6. Servermodul nach einem der vorangehenden Ansprüche,
dadurch gekennzeichnet,
daß das Routing-Servermodul die aktuelle Routing-Tabelle an
die übrigen Servermodule über die Schaltschnittstelle (7)
überträgt.

10

7. Servermodul nach einem der vorangehenden Ansprüche,
dadurch gekennzeichnet,
daß das Routing-Servermodul Daten über die Auslastung der Da-
tenverarbeitungseinheiten aller Servermodule des modular auf-
15 gebauten Servers (1) sammelt und auswertet.

8. Servermodul nach einem der vorangehenden Ansprüche,
dadurch gekennzeichnet,
daß das Routing-Servermodul die Routing-Tabelle in Abhängig-
20 keit von den ausgewerteten Auslastungsdaten, den zugewiesenen
Applikationstypen der übrigen Servermodule sowie von Priori-
tätsinformationsdaten der übertragenen Datenpakete aktuali-
siert.

25 9. Servermodul nach einem der vorangehenden Ansprüche,
dadurch gekennzeichnet,
daß bei einer zu hohen Auslastung der Datenverarbeitungsein-
heit (17) des Servermoduls Datenverarbeitungsprozesse, die
innerhalb des Servermoduls ablaufen, auf Datenverarbeitungs-
30 einheiten von anderen Servermodulen übertragen werden.

Zusammenfassung

Servermodul für modular aufgebauten Server

- 5 Servermodul für modular aufgebauten Server (1) mit
mindestens einer Datenverarbeitungseinheit (17) zur Datenver-
arbeitung von Datenpaketen, mindestens einer adressierbaren
Kommunikationsschnittstelle (19) zum Anschluß des Servermo-
duls (2, 3, 4, 5) an ein externes Netzwerk (13), über das die
10 Datenpakete übertragen werden, einer Schaltschnittstelle (20)
zum Anschluß des Servermoduls (2, 3, 4, 5) an eine Schaltein-
richtung (7) des modular aufgebauten Servers (1); und mit ei-
ner Routing-Berechnungseinheit (18) zur Berechnung einer Ser-
vermodul-Adresse mittels einer Routing-Tabelle in Abhängig-
15 keit von der Auslastung der Datenverarbeitungseinheiten aller
Servermodule (2, 3, 4, 5) des modular aufgebauten
Servers (1).

Fig. 2

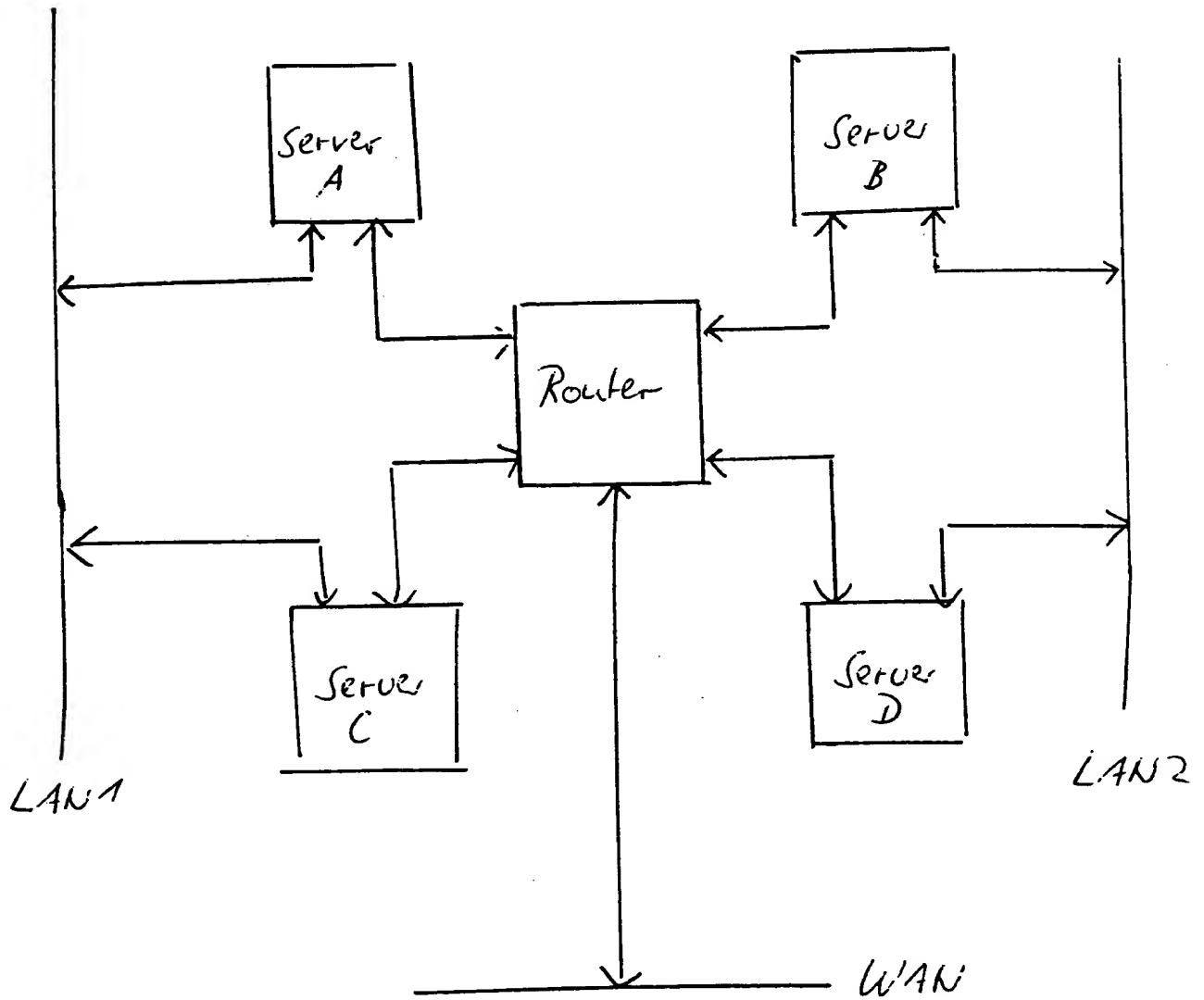


Fig. 1

StdT

2/3

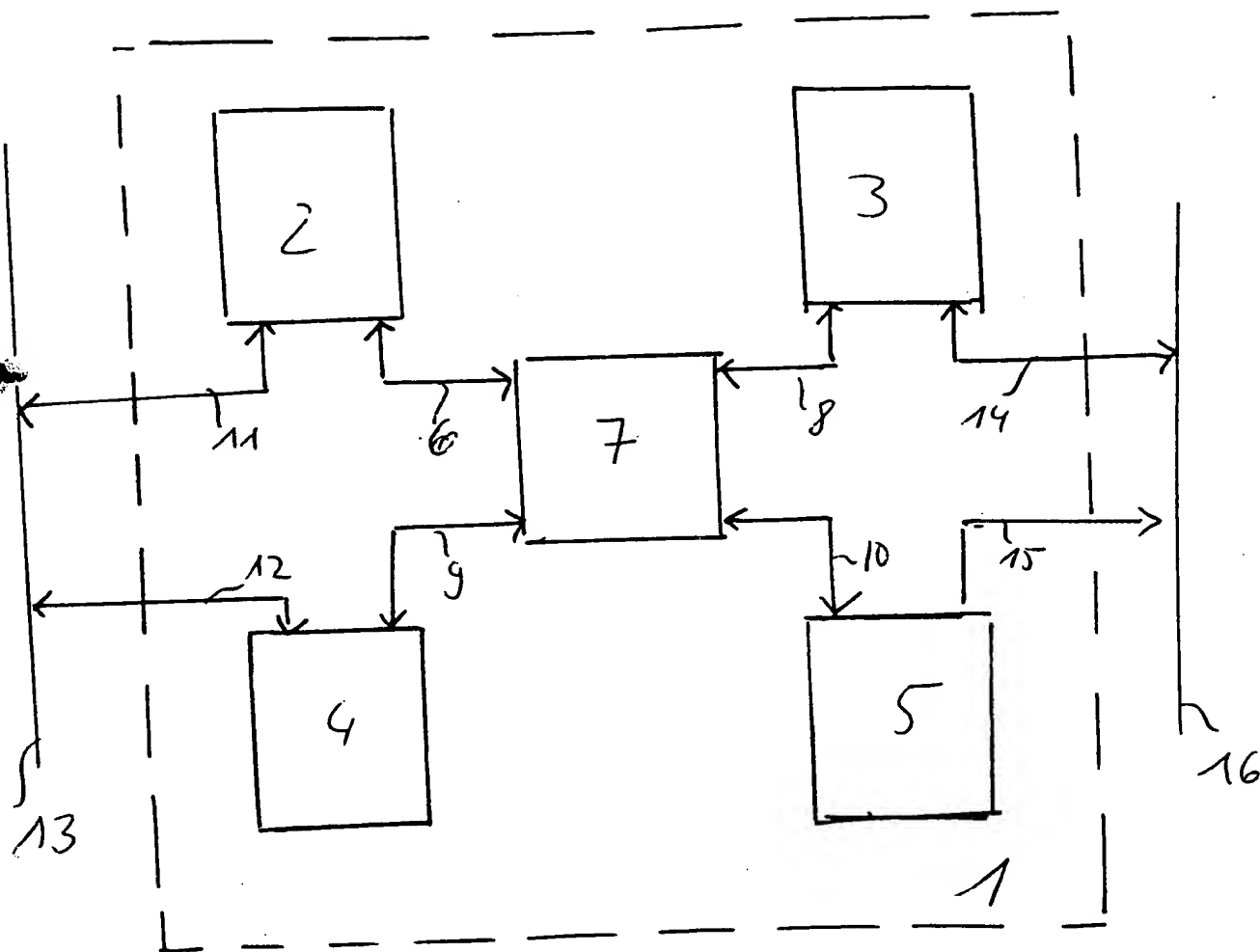


Fig. 2

3/3

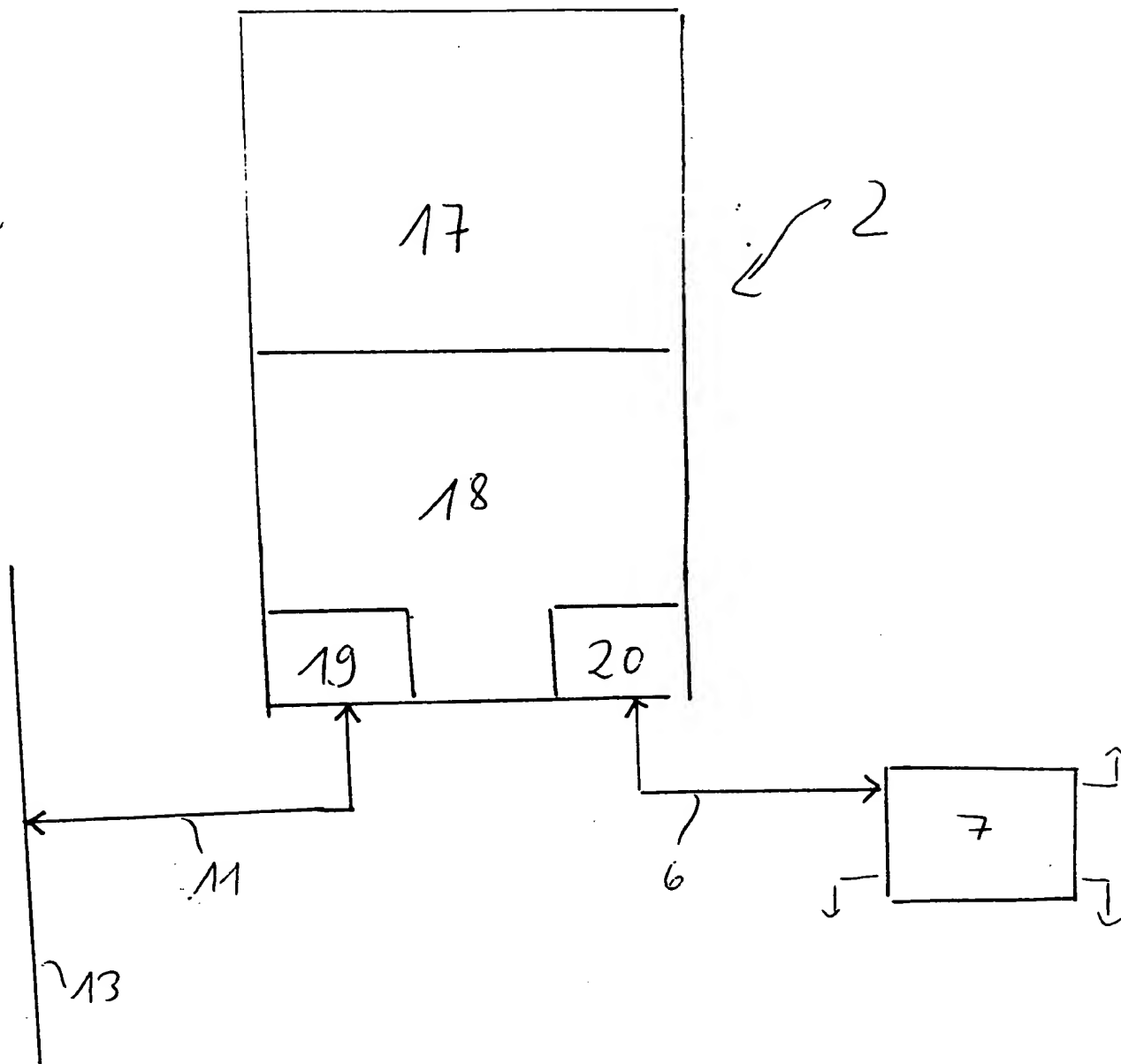


Fig. 3